PERFORMANCE OF UMTS PACKET SERVICE OVER DEDICATED CHANNELS (DCH)

Antonio Capone, Matteo Cesana
Politecnico di Milano, Dipartimento di Elettronica e Informazione, Italy
e-mail: {capone, cesana}@elet.polimi.it

Abstract—
The provision of high speed packet data services in an efficient way is probably the most important challenge for UMTS since this can give an advantage with respect to second generation systems. The efficiency mainly depends on the radio interface and since it is characterized by a great flexibility it is of utmost importance to investigate the effect of different configurations on the system performance. In this paper we evaluate the performance of packet data services over downlink dedicated channels (DCH) by means of detailed simulations. Due to the traffic variability, ARQ mechanism and closed loop power control the system behavior is quite complex and not so easily predictable as with constant rate services such as voice. We study the effect of several parameters (spreading factor, code rate, channel setup delay, etc.) on the system capacity by means of the delay-throughput curves. We show that the setting of parameters may be critical for system capacity and stability. For this reason we propose a flow control mechanism which is able to guarantee stability and to make the system capacity almost independent of channel rates.

I. INTRODUCTION

The success of mobile cellular systems has completely changed the way users access telecommunication services. However, even if cellular systems are now widely adopted for telephone services, the number of users applying for data services remains very low. This is due to the second generation systems technological limitations of bandwidth and service flexibility. As an example, the GSM (Global System for Mobile communications) have only provided until now low rate circuit switched data services which are not suitable even for web or e-mail applications due to high latency times and high costs.

To fill the gap between user needs and services offer, second generation systems are being enhanced to include packet data services, such as GPRS (General Packet Radio Service), which allow a more flexible use of radio resources and higher peak rates [1]. The commercial introduction of these new services (usually referred to as 2.5 generation services) is now coming while the standardization process of third generation systems is being completed. It is therefore evident that this poses new challenges to third generation systems since they must prove their ability to provide access for a wide range of multimedia applications and services in an efficient and cost effective way.

Universal Mobile Telecommunications System (UMTS) [2], [3], [4] is the third generation mobile communication system standardized by 3GPP, the Third Generation Partnership Project, and is also considered by ITU (International Telecommunication Union) among the standards for the IMT-2000 (International Mobile Telephone standard 2000) family [5], [6]. One of the two access schemes to be used in the assigned spectrum is based on W-CDMA (Wideband Code Division Multiple Access) and frequency division duplexing (FDD) [7].

One of the main advantage of UMTS and all the 3G systems with respect to the second generation is the capability of providing radio access to multimedia services. The traffic being transferred within 3G mobile networks can be composed by different information flows with various constraints on the required QoS (bit rate, delays, etc...). In order to do the job, the radio interface of UMTS is characterized by great flexibility and a variety of different physical and logical channel types. For instance, several user rates and channel code protections are available by choosing suitable parameters, such as spreading factors, FEC (Forward Error Correction) rates and ARQ (Automatic Repeat reQuest) schemes. This approach is quite different from the one adopted by second generation systems where a small set of services can be implemented by vendors and provided by operators. If from one side this added flexibility is an advantage of UMTS, from the other it makes the task of real services implementation more complex. In fact, the complexity of detailed services definition and system parameters optimization has been moved out of specifications and let UMTS vendors and operators.

In this scenario the optimization of radio interface parameters is of utmost importance both in circuit switched service and in packet switched ones. Even if CDMA systems have been widely analyzed in the literature, to the best of our knowledge, the effects of the different parameters settings on the performance of UMTS data services has not been thoroughly investigated. In [8] we studied the performance of the UMTS Downlink Shared Channel (DSCH) where packet flows are time multiplexed and spreading codes are used basically to fight interference due to transmissions in other cells (inter-cell interference). In this paper we evaluate the performance of packet data services over downlink dedicated channels (DCH). These channels are the same adopted for circuit switched services such as voice, but with packet service their behavior changes significantly due to varying interference.

The paper is organized as follows. In Section 2 we give a short overview of the basic characteristics of UMTS radio interface, and we describe the model used for simulations. In

Section 3 we present and discuss the results obtained. Finally, Section 4 includes our concluding remarks.

II. Radio Access Overview and Simulation Model

A. Radio Access Overview

UMTS specifications [9] considers two access schemes, W-CDMA (Wideband CDMA) and TD-CDMA (Time Division and Code Division Multiple Access), to be used respectively in the paired part of the spectrum assigned to UMTS, 60 MHz from 1920 to 1980 MHz (uplink) and 60 MHz from 2110 to 2170 MHz (downlink), and in the unpaired part, 35 MHz from 1900 to 1920 MHz and from 2010 to 2025 MHz, respectively. The W-CDMA scheme adopts a chip rate of $3,840$ Mchip/s. It presents a carrier separation of 5 MHz, so that up to 12 carriers can be defined in the available bandwidth. For the downlink direction a QPSK modulation is adopted after spreading, while for the uplink direction the in-phase and quadrature channels are used to transmit two BPSK flows which may have different spreading codes [10].

At the air interface, the physical layer offers a transport service to higher layers through physical channels. The upper layer information is first protected by the physical layer using FEC (Forward Error Correction) codes [11] and then it is spread and modulated with a constant chip rate. The FEC scheme uses convolutional codes with rate $1/3$ or $1/2$, or a turbo code with rate $1/3$. Different rates can be obtained using the rate matching process which can increase the code rate by means of puncturing. The spreading process is based on two codes, namely the spreading code and the scrambling code. The spreading code increases the flow bit rate to the chip rate of the air interface according to the Spreading Factor (SF). Different values of SF ranging from 4 to 512 are available and they are obtained using a tree of orthogonal codes. The tree has the characteristic that two codes, even with different SF, are orthogonal if they are located in different branch of the tree (OVSF). Multiple trees can be generated using a scrambling code which varies the order of chips. The channels transmitted by the same station (base or mobile) can use codes in the same tree so that the mutual interference is greatly reduced, while channels transmitted by different stations should use different scrambling codes so that the mutual interference is almost independent of the delay offset at the receiver due to different propagation paths.

Physical channels are defined by the associated spreading and scrambling codes. The bit flow is divided into time-slots, 666 μs long. During a time-slot both physical control bits and data bits can be transmitted, and while the number of chips is fixed, the total number of bits depends on the SF. The minimum transmission unit offered by the physical layer to the upper layers is the Time Transmission Interval (TTI), also called frame in the following, and is composed of multiples of 15 slots. The transport services provided by the physical layer to the upper layers are based on transport channels which are mapped into the physical channels [12].

Transport channels are divided into dedicated channels, which can be assigned and then used only for transmissions to/from a single mobile terminal (MT) at a time, and common channels which are time shared by different MTs. Dedicated Channels (DCH) are used to transmit user and control information in the uplink and downlink direction and are devoted to the connection between a single mobile station and the UTRA Network (UTRAN).

Power control procedures are defined at the physical layer to adjust transmission power of physical channels according to interference and propagation conditions. The power control exerted on the DCH is based on a closed-loop signaling, outlined in the following. The TPC (Transmit Power Control) symbols are transmitted in each slot and carry a command for increasing or decreasing the transmission power in each direction. The power step is fixed and usually is equal to 1 dB. On the receiving side, if the estimated SIR (Signal-to-Interference Ratio) after despreading is lower than a SIR target value, an increase command is sent in the subsequent slot. A decrease command is sent otherwise. The SIR target value is controlled by an outer control loop which is based on the quality of the decoded bit flow.

In the user plane on top of the physical layer, the link layer is split into the MAC (Medium Access Control), the RLC (Radio Link Control) and PDCP (Packet Data Convergence Protocol) [13]. The MAC layer [14] provides logical channels to the RLC and maps logical channels into transport channels. On common channels, the MAC provides addressing of user equipments and scheduling of PDUs (Packet Data Units). The RLC layer [15] can offer an acknowledged or unacknowledged data transfer mode.

B. Simulation Model

In order to help reader's comprehension of the results that are being presented, we give here a brief outline on the simulation model we adopted. A detailed description of the simulated scenarios is reported in [8].

We have considered 49 cells with radius equal to 300 m, organized in a wrap–around domain to avoid border effects in the interference calculation. Each cell has an omni-directional antenna with unit gain located at the center.

The received power P_r of the generic downlink transmission is given by [17]: $P_r = P_t 10^{\frac{\epsilon}{10}} L$, where P_t is the transmitted power, L is the path loss, $10^{\frac{\epsilon}{10}}$ accounts for the loss due to slow shadowing, being ϵ a normal variate with zero mean and σ^2 variance. At this stage of analysis we have considered a macro-cellular environment, for which the path loss L is expressed as: $10 \log L = -(128.1 + 37.6 \log r)(dB)$, where r (in kilometers) represents the distance between the mobile and the base station. Furthermore, we assume shadowing standard deviation equals to 10 dB. The generic user is assigned to the BS with the minimum attenuation.

At the receiving side, the SIR is evaluated, for each transmission, as

$$SIR = \frac{P_r \times SF}{\alpha I_{intra} + I_{inter} + P_N} \qquad (1)$$

where SF is the spreading factor of the physical connection, P_N is the thermal noise assumed equal to -99 dBm, I_{inter} is the sum of the signal powers received from the other cells, I_{intra} is the sum of the signal powers received from other users in the same cell, and α is the loss-of-orthogonality factor due to the multipath that, according to [16], is assumed equal to 0.4.

The SIR calculated value is used to test correctness of the transmission. Our simulator assumes an ideal ARQ procedure, i.e. the transmitted block is kept in the transmitting queue in case of error and is canceled otherwise. After 10 failed transmissions the block is dropped and the user is declared in outage.

The dedicated channels are subject to a typical closed loop power control procedure. The MT request the BS for a transmitting power update in order to follow the SIR fluctuations.

Physical power constraint are also added as specified in [16]. The power transmitted on each downlink DPCH can not exceed the value 30 dBm, whereas the overall power transmitted by a base station is limited to 43 dBm.

The birth of new user in the system is modeled with a Poisson point process of intensity λ [17]. Each new user request the download of a web–page which is modeled by a flow of packets whose number is geometrically distributed with mean $N_p = 25$. The packet length is negative exponentially distributed with mean 3840 bits. Packets composing the web–page arrive at the Base Station buffers according to a Poisson point process whose intensity depends on the traffic source speed. Before transmission, the physical layer at UTRAN side adds redundancy bits due to the Cyclic Redundancy Check procedure and to the coding scheme adopted [11].

Our simulator adds the parity bits required by Convolutional Codes, with 256 states, Constraint Length $K = 9$ and optimal puncturing, whose Bit Error Probability (BER), obtained through link level simulations [18].

III. Simulation Results

All the presented results have been obtained running steady-state simulations 600 seconds long. The first 100 seconds are used as warm–up time, that is to say the statistical results collected during this period are neglected. The remaining 500 seconds are divided into 5 simulation runs. During each run the results are collected and used to calculate one sample of each statistical quantity used for evaluation. The output results have been tested according to the t-student statistical test. For all the measures reported in the following (throughput, BLER, etc) the confidence interval is under the 5%, given a confidence level of 95%.

A. Delay–Throughput performances of the Dedicated Channel Service

As outlined in section II-A, UMTS allows to set transport channel parameters in a flexible way in order to optimize system performance according to traffic characteristics and interference conditions. In this subsection we study the performance of the DCH service when varying the spreading factor of the physical channel to be used. To characterize the performance we consider the curves of the average packet-delivery delay versus the achieved throughput.

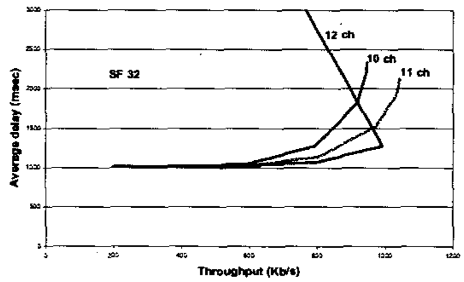

Fig. 1. Average delay vs. Throughput for $SF = 32$ and different number of DCHs.

As in the shared service [8], the optimal value of SIR target comes from a trade–off choice. If a too small SIR target is chosen too many errors occur since the SIR fluctuations around the target value often drive the system in a condition where the code protection is useless. On the other side, with a high SIR target the power requirement increases and too many transmissions tend to be driven in saturation. All the results that are being presented have been obtained with optimized SIR target values with respect to the maximal throughput and BLER. The optimal SIR target value has been found to be 4 dB for any spreading factor configuration considered.

Figure 1 shows the average delay vs throughput curves when the service uses 10, 11 and 12 SF= 32 channels per cell and the information is protected with a R= $1/2$ convolutional code. This configuration provides a neat bit rate entering the physical layer of 105.9 kb/s. The best performance is achieved with 11 channels per cell, with a maximal throughput of 1050 kb/s.

We observe that, if the contemporary use of 12 dedicated channels per cell is allowed, the system is driven to instability and the average delay vs. throughput curve bends backwards. Naturally, when we increase the number of dedicated channels in each cell we increase also the interference. The power control procedure tries to counteract such an interference increase by requesting base stations to transmit more power. However, if a new equilibrium point (a power levels assignment which gives all SIR values equal to the target one) does not exist the closed loop control mechanism increases all transmission powers until some of them reach the upper limit seen in section II-B. In these conditions most of the connections cannot reach the target SIR value and, as a result, the fraction of erred transmission increases and the throughput decreases. Moreover, due to the ARQ mechanism which tries to recover lost packets, the traffic on the channel becomes higher and the interference further increases.

This behavior shows that also with packet data service we

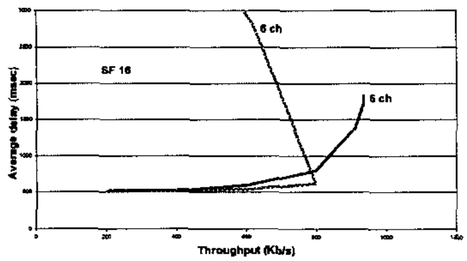

Fig. 2. Average delay vs. Throughput for $SF = 16$ and different number of DCHs.

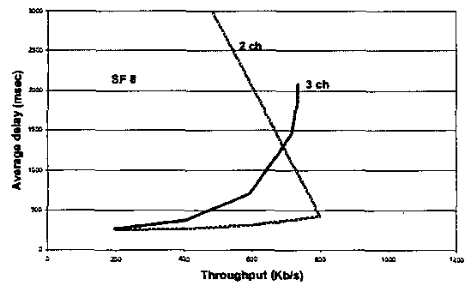

Fig. 3. Average delay vs. Throughput for $SF = 8$ and different number of DCHs.

need to limit the number of resources available for the service, since there exists a limit on the interference which is tolerable by the system. If not, the power control procedure is not effective in facing interference at high loads and the system efficiency may become very poor. Since we are dealing with dedicated channels which are assigned to single users through a set-up procedure, we can say that such a limitation is a call admission control even if from the service point of view there is no call (the set-up and tear-down should be transparent to packet service users). The limitation of the number of contemporary active dedicated channels per cell is a simple control that can be adopted, even if more flexible mechanisms (out of the scope of this paper) could be designed [19].

A similar behavior has been observed using different spreading factor services. Figures 2 and 3 shows the average delay versus the throughput when using respectively 3, 4 SF= 8 channels and 5, 6 SF= 16 channels per cell. Also in this two cases, there is a maximum level of interference affordable by the system, and a maximum number of channels to be fixed to prevent instability, which is respectively 3 in the SF= 8 configuration and 5 in the SF= 16 one.

Figure 4 resumes the results of figures 1, 2 and 3, showing the best delay-throughput curves obtained for the three configuration of the dedicated service (SF= $8, 16, 32$). The configuration with SF= 32 channels is shown to be optimal from the maximal throughput point of view, but it presents high packet delivery delays even for low throughput values. On the other hand, the configuration with SF= 8 achieves lower throughput values but grants limited delay for lower

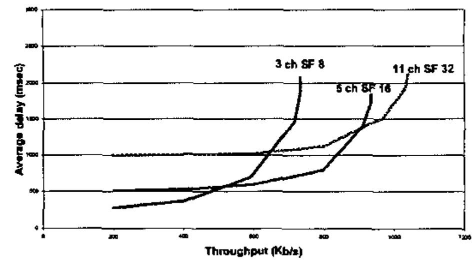

Fig. 4. Average delay vs. Throughput. Comparison of the best curves for the $SF = 8, 16, 32$ configurations.

load values. This result confirms that, with low rate channels, the power control mechanism can better track SIR target due to the reduced interference variance, even if the average level increases due to the added intra-cell interference.

B. Flow Control Mechanism

Instability has been proved to occur when the mean interference level gets high and the power control drives in saturation the transmitted power. In such a condition the power control cannot prevent transmission errors, since it is forced to request for more power that is not available. Limiting the number of channels per cell is a first possible approach to prevent performance degradation at high loads. However, since this limit must be set in advance, it is necessary to take into account the worst possible scenarios so assuming a conservative constraint.

However, since we are considering packet data service, an additional flow control (FC) mechanism can be adopted in order to dynamically adjust the load on the active channels, and therefore the average interference generated, according to traffic and propagation conditions. Such a FC scheme must be able to limit the transmission rate only when necessary reducing the intensity and frequency of interference peaks.

We have proposed and implemented in our simulator a dynamic FC mechanism which is based on a feedback provided by mobile terminals. The basic idea is to reduce the transmission rate on the channel when one or more consecutive transmissions fails. However, since a transmission can fail also due to interference variations and not due to a real traffic congestion (we could say interference congestion), the mechanism is triggered only when a transmission performed at the maximum powers. More in details, a flow-control timer is started so that transmissions are inhibited for a random number of frames (10 ms long) uniformly distributed in the interval $(1, n)$, where n is the minimum between the number of consecutive wrong transmissions performed at maximum power and $n_{max} = 10$. In such a way, we control the traffic on the channels (G), and consequently we limit the mean interference level. In order to let available the information on the transmission result at the transmitting end, we adopt one of the FBI (Feedback Information) bits defined in the uplink slot format of the DPCH [12].

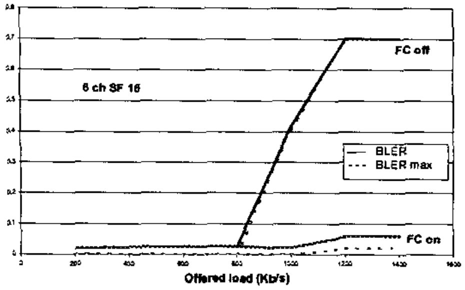

Fig. 5. BLER vs. Offered load with $SF = 16$. Comparison of the cases with and without the FC mechanism.

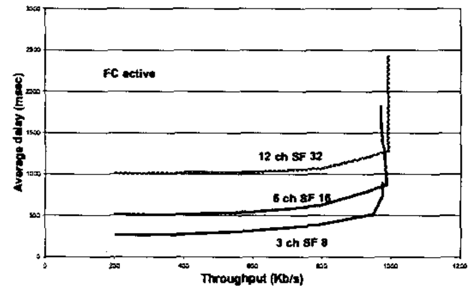

Fig. 6. Average delay vs. Throughput with the FC mechanism. Comparison of the best curves for the $SF = 8, 16, 32$ configurations.

To show the effect of the proposed FC mechanism we have reported in Figure 5 the Block Error Rate (BLER) and the Block Error Rate when the transmission is at maximum power (BLER max) versus the offered load in the two cases with and without the FC mechanism. The physical layer configuration used is the one with 6 SF= 16 channels. Without FC, the BLER is very high and almost equal to the BLER max (about 0.7 at high loads). This show that most of the errors occur when the power control cannot reach an equilibrium point and brings the power at its maximum value. In contrast, in the case with the FC mechanism the BLER is lower (0.1 at high loads) and, furthermore, only a relatively small fraction of errors are due to power limitations. A similar behavior has been observed with the other physical layer configurations (SF= 8, 32).

The throughput vs average delay curves obtained with the FC mechanism have been summarized Figure 6 which reports for SF= 8, 16, 32 only the cases with the highest maximum throughput. In contrast with the results obtained without FC (see Figure 4), the maximum throughput achieved with the three SF values are almost the same. We observed that at high loads the FC mechanism forces the system with all the spreading factor configurations to work with the same average interference levels by modulating the traffic on the channels (G). This allows also configurations with high rate channels (small SF) to reach a high throughput while maintaining a low delay at high loads. Moreover, in all configurations the system is able to handle one more channel than the corresponding cases without the FC mechanisms.

IV. Conclusions

In this paper we have presented the preliminary results on the performance of the UMTS dedicated channel service when transporting packet switched traffic.

We have analyzed the behavior of the system when varying the speed of the physical dedicated channel to be used (SF=8, 16, 32). For all the three spreading factor configurations, we have proved that there exists a limit on the mean interference which can be tolerate by the system. Beyond this limit, the closed loop power control can't provide the SIR target to many active connections and the system is driven to instability.

In order to prevent instability, we have proposed a flow control algorithm, which dynamically controls traffic on the channels and forces the system to work with a tolerable average interference level.

References

[1] R. Kalden, I. Meirick, M. Meyer, Wireless Internet access based on GPRS, IEEE Personal Communications, vol. 7, no. 2, April 2000, pp. 8-18.

[2] A. Samukic, *UMTS universal mobile telecommunications system: development of standards for the third generation*, IEEE Transactions on Vehicular Technology, vol. 47, no. 4, Nov. 1998, pp. 1099-1104.

[3] K.W. Richardson, *UMTS overview*, Electronics & Communication Engineering Journal, vol. 12, no. 3, June 2000, pp. 93-100.

[4] M. Gallagher, W. Webb, *UMTS the next generation of mobile radio*, IEE Review, vol. 45, no. 2, March 1999, pp. 59-63.

[5] T. Ojanpera, R. Prasad, *An overview of air interface multiple access for IMT-2000/UMTS*, IEEE Communications Magazine, vol. 36, no. 9, Sept. 1998, pp. 82-95.

[6] D. O'Mahony, *UMTS: the fusion of fixed and mobile networking*, IEEE Internet Computing, vol. 2, no. 1, Jan.-Feb. 1998, pp. 49-56.

[7] 3rd Generation Partnership Project, Physical layer - General description, 3G TS 25.201, June 2000.

[8] F. Borgonovo, A. Capone, M. Cesana, L, Fratta, *Packet service in UMTS: delay-throughput performance of the downlink shared channel (DSCH)*, Journal of Computer Networks, vol. 38, no. 1, Jan. 2002, pp. 43-59.

[9] UMTS Technical Specifications, public documents available at http://www.3gpp.org.

[10] 3rd Generation Partnership Project, Spreading and modulation (FDD), 3G TS 25.213, June 2000.

[11] 3rd Generation Partnership Project, Multiplexing and channel coding (FDD), 3G TS 25.212, June 2000.

[12] 3rd Generation Partnership Project, Physical channels and mapping of transport channels onto physical channels (FDD), 3G TS 25.211, June 2000.

[13] 3rd Generation Partnership Project, Radio Interface Protocol Architecture, 3G TS 25.301, June 2000.

[14] 3rd Generation Partnership Project, MAC protocol specification, 3G TS 25.321, June 2000.

[15] 3rd Generation Partnership Project, RLC Protocol Specification, 3G TS 25.322, June 2000.

[16] 3rd Generation Partnership Project, RF system scenarios, 3G TR 25.942, Dic. 1999.

[17] UMTS 30.03, Annex B: Test environments and deployment models, TR 101 1112 v.3.2.0, April 1998.

[18] A. Bellini, M. Ferrari, Personal communication, Politecnico di Milano, 2000.

[19] A. Capone, S. Redana, Call Admission Control Schemes for UMTS, IEEE VTC 2001 Fall, Atlantic City, NJ, Oct. 2001.